

# Distribution of work in isothermal non-equilibrium processes

Thomas Speck and Udo Seifert

*II. Institut für Theoretische Physik, Universität Stuttgart, 70550 Stuttgart, Germany*

Diffusive motion in an externally driven potential is considered. It is shown that the distribution of work required to drive the system from an initial equilibrium state to another is Gaussian for slow but finite driving. Our result is obtained by projection method techniques exploiting a small parameter defined as the switching rate between the two states of the system. The exact solution for a simple model system shows that such an expansion may fail in higher orders, since the mean and the variance following from the exact distribution show non-analytic behavior.

PACS numbers: 05.40.-a, 05.70.Ln

In macroscopic thermodynamics, the work  $W$  spent in changing the state of a system at constant temperature  $T$  obeys

$$W \geq \Delta F, \quad (1)$$

which is one version of the second law where  $\Delta F$  is the difference in free energy of the final and the initial equilibrium state. As the system gets smaller, thermal fluctuations play an increasingly relevant role. Hence, this work acquires a stochastic contribution, i.e., the work follows a distribution function  $P(W)$ . The shape of this function depends on how the system is driven. If this change is induced by the time variation of an external control parameter  $\lambda(t)$ , the distribution  $P(W)$  becomes a functional of  $\lambda(t)$ .

Such distributions have recently become accessible experimentally for systems with only a few degrees of freedom diffusing in a thermal environment under the influence of an externally controlled potential. Paradigmatic examples include dragging a colloid particle by an optical tweezer through a viscous fluid [1, 2] and the forced unfolding of RNA hairpins [3]. In both cases some realizations of the process show  $W < \Delta F$ . Slightly overstated, such findings have been called violations of the second law [1]. In a more conservative interpretation of the second law for such mesoscopic systems, the average work should and does still obey

$$\overline{W} \equiv \int_{-\infty}^{+\infty} dW P(W)W \geq \Delta F. \quad (2)$$

Obviously, the distribution  $P(W)$  is of paramount importance for a better understanding of isothermal stochastic dynamics. Exact statements, however, about  $P(W)$  are scarce. In 1997, Jarzynski has shown under rather mild assumptions that the distribution  $P(W)$  obeys an integral constraint

$$\int_{-\infty}^{+\infty} dW P(W)e^{-\beta W} = e^{-\beta \Delta F} \quad (3)$$

for any external protocol  $\lambda(t)$  [4, 5, 6]. Here,  $\beta \equiv 1/k_B T$  with Boltzmann's constant  $k_B$ . Since this remarkable

relation allows one to extract equilibrium free energy differences from measuring or calculating the work distribution in non-equilibrium experiments or simulations, it has found widespread applications recently [3, 7, 8, 9, 10, 11]. The statistical and convergence properties of this non-linear average deserve particular attention [12, 13, 14].

For time-dependent quadratic potentials, i.e. linear stochastic equations of motion,  $P(W)$  can easily derived to be Gaussian [15]. Jarzynski's relation then implies that the mean  $\overline{W}$  and the variance  $\sigma^2$  are necessarily related by [4]

$$\overline{W} = \Delta F + \beta\sigma^2/2. \quad (4)$$

For these potentials, the Gaussian nature holds independently on the speed of the driving, i.e. independently of how far the system is from equilibrium.

The purpose of this paper is to add a third general statement about  $P(W)$  to this list of exact results. We will show that this distribution becomes a Gaussian for slow but finite driving even if the equations of motion are non-linear. Since our approach is constructive, it yields an explicit algorithm of how to obtain the mean  $\overline{W}$  and the variance  $\sigma^2$  of this Gaussian distribution. In the quasi-static limit of infinitely slow external manipulation, this Gaussian reduces to  $P(W) = \delta(W - \Delta F)$ .

A Gaussian character of the distribution  $P(W)$  near equilibrium seems to be expected or taken for granted in the recent literature [3, 4, 10, 13, 16]. Closer scrutiny of the references usually cited for this assumption – if any are cited at all –, however, reveals that they do not provide an explicit proof of this statement. The often cited papers by Hermans [17] and Wood et al. [18] explicitly assume a Gaussian shape. Alluding in a more general way to an Onsager-Machlup functional [19] also fails, since this Gaussian functional is derived for linear stochastic equations of motion. The presumably most promising case to date in favor of a Gaussian distribution suggests to invoke the central limit theorem for the increments of work [20]. However, without translating this proposal into a definite calculation, which seems to be non-trivial for time-dependent potentials, this argument is not a clear-cut proof yet, let alone does it give an

expression for mean and variance of this putative Gaussian.

Based on this unsatisfying state of affairs concerning such a fundamental issue, we believe that a constructive derivation of the Gaussian nature of this distribution for finite but slow driving is indeed called for as a step towards a comprehensive theory of isothermal stochastic dynamics.

For the derivation we consider a finite classical system coupled to a heat reservoir of constant temperature. Let then  $\mathbf{x} = (x_0, \dots, x_n)$  be the state of the system with energy  $V_\lambda(\mathbf{x})$  where  $\lambda$  is an externally controlled parameter. The stochastic dynamics is governed by the Langevin equations [21]

$$\dot{x}_i = -\mu_{ij} \frac{\partial V_\lambda}{\partial x_j} + \eta_i(t) \quad (5)$$

where  $\mu_{ij}$  are the mobility coefficients [24] and  $\eta_i(t)$  is the thermal noise representing the heat bath with

$$\langle \eta_i(t) \rangle = 0 \text{ and } \langle \eta_i(t) \eta_j(t') \rangle = \frac{2}{\beta} \mu_{ij} \delta(t - t'), \quad (6)$$

where  $\langle \dots \rangle$  denotes the ensemble average. We describe the continuous process of switching the system from an initial state ( $\lambda(t=0) = 0$ ) to a final state ( $\lambda(t=t_s) = 1$ ) by a protocol  $\lambda(t)$ , over a total switching time  $t_s$ . Without loss of generality, we set

$$\lambda(t) \equiv t/t_s \quad (7)$$

and hence a constant switching rate  $\dot{\lambda} = t_s^{-1}$  [4].

We now consider an ensemble of infinitely many realizations of this Markov process, each evolving stochastically according to eq. (5). The normalized distribution of this ensemble in phase space  $f(\mathbf{x}, t)$  obeys a Fokker-Planck equation [21]

$$\partial_t f = \hat{L}_\lambda f \text{ with } \hat{L}_\lambda \equiv \frac{\partial}{\partial x_i} \mu_{ij} \left[ \frac{\partial V_\lambda}{\partial x_j} + \frac{1}{\beta} \frac{\partial}{\partial x_j} \right] \quad (8)$$

equivalent to the Langevin equations (5). This introduces the (through  $\lambda$  time-dependent) Fokker-Planck operator  $\hat{L}_\lambda$ . The stationary solution of eq. (8) for fixed  $\lambda$  is the equilibrium distribution

$$f_\lambda^{\text{eq}}(\mathbf{x}) \equiv e^{-\beta V_\lambda(\mathbf{x})} / \int d\mathbf{x}' e^{-\beta V_\lambda(\mathbf{x}')}. \quad (9)$$

The total work performed along one particular trajectory  $\mathbf{x}(t)$  up to time  $t$  is the time integral [4, 22]

$$W[\mathbf{x}(t), t] \equiv \int_0^t dt' \dot{\lambda} \frac{\partial V_\lambda}{\partial \lambda}(\mathbf{x}(t')). \quad (10)$$

We can now compose a combined stochastic process consisting of  $\{\mathbf{x}, W\}$  as [15]

$$\dot{x}_i = -\mu_{ij} \frac{\partial V_\lambda}{\partial x_j} + \eta_i(t) \quad (11)$$

$$\dot{W} = \dot{\lambda} \frac{\partial V_\lambda}{\partial \lambda}. \quad (12)$$

Note that the equation of motion for  $\dot{W}$  does not have an independent noise but is stochastic through the  $\mathbf{x}$ -dependence of  $V_\lambda$ . The joint probability distribution function  $p(\mathbf{x}, W, t)$  then obeys a Fokker-Planck equation

$$\partial_t p = [\hat{L}_\lambda + \dot{\lambda} \hat{L}_\lambda^W] p \quad (13)$$

where

$$\hat{L}_\lambda^W \equiv -\frac{\partial V_\lambda}{\partial \lambda} \frac{\partial}{\partial W} \quad (14)$$

represents a drift term of the work. Eq. (13) is exact, no matter how far the system is driven out of equilibrium. The reduced probability distribution of the work  $P(W, t)$  can be obtained by integrating out  $\mathbf{x}$  as

$$P(W, t) = \int d\mathbf{x} p(\mathbf{x}, W, t). \quad (15)$$

Since we start the process out of thermal equilibrium, the  $\mathbf{x}$  are initially distributed according to the canonical distribution and therefore the initial condition is

$$p(\mathbf{x}, W, 0) = f_0^{\text{eq}}(\mathbf{x}) \delta(W). \quad (16)$$

As our main theoretical tool we introduce a projector  $\hat{\Pi}_\lambda$  acting on a function  $\phi(\mathbf{x}, W, t)$  such that

$$\hat{\Pi}_\lambda \phi \equiv f_\lambda^{\text{eq}} \int d\mathbf{x}' \phi(\mathbf{x}', W, t). \quad (17)$$

Note that

$$\hat{\Pi}_\lambda \hat{\Pi}_\lambda \phi = \hat{\Pi}_\lambda \hat{L}_\lambda \phi = 0. \quad (18)$$

The first statement is evident from definition (17) and the fact that  $f_\lambda^{\text{eq}}$  is in the null space of  $\hat{L}_\lambda$ . The second conclusion follows when  $\phi$  is expanded in terms of eigenfunctions of  $\hat{L}_\lambda$ . Then the Fokker-Planck operator  $\hat{L}_\lambda$  cancels the eigenfunction to eigenvalue 0, which in fact is  $f_\lambda^{\text{eq}}$ , whereas the projector annihilates all other eigenfunctions corresponding to higher eigenvalues.

The other important property of the projector  $\hat{\Pi}_\lambda$ , which distinguishes it from the usual application to the adiabatic elimination of fast variables [21], is that it does not commute with the time derivative but rather leads to

$$[\partial_t, \hat{\Pi}_\lambda] \equiv \partial_t \hat{\Pi}_\lambda - \hat{\Pi}_\lambda \partial_t = -\dot{\lambda} \beta S_\lambda \hat{\Pi}_\lambda, \quad (19)$$

where we define

$$S_\lambda \equiv \frac{\partial V_\lambda}{\partial \lambda} - \left\langle \frac{\partial V_\lambda}{\partial \lambda} \right\rangle_\lambda. \quad (20)$$

The equilibrium ensemble average  $\langle \dots \rangle_\lambda$  is defined as

$$\langle \phi \rangle_\lambda \equiv \int d\mathbf{x} f_\lambda^{\text{eq}}(\mathbf{x}) \phi(\mathbf{x}, t). \quad (21)$$

We can now expand the joint probability  $p(\mathbf{x}, W, t)$  for small  $\dot{\lambda}$ , which corresponds to a separation of time scales [23]. The slow time scale is  $\lambda = t/t_s$ . The fast time scale, which we do not need explicitly, is determined by the intrinsic relaxation processes. As the time derivative transforms according to  $\partial_t \rightarrow \dot{\lambda} \partial_\lambda$ , switching to the slow time scale eq. (13) becomes

$$\partial_\lambda p = \left[ \dot{\lambda}^{-1} \hat{L}_\lambda + \hat{L}_\lambda^W \right] p. \quad (22)$$

By using the projector  $\hat{\Pi}_\lambda$  we decompose the distribution function  $p = p_0 + p_1$  into

$$p_0(\mathbf{x}, W, \lambda) \equiv \hat{\Pi}_\lambda p = f_\lambda^{\text{eq}}(\mathbf{x}) P(W, \lambda) \quad (23)$$

and

$$p_1(\mathbf{x}, W, \lambda) \equiv (1 - \hat{\Pi}_\lambda)p. \quad (24)$$

We apply  $\hat{\Pi}_\lambda$ , respectively  $(1 - \hat{\Pi}_\lambda)$ , to eq. (22) and keep in mind both eq. (18) and the commutator (19). We finally get the two coupled differential equations

$$\partial_\lambda p_0 = \hat{A}_\lambda^0 p_0 + \hat{A}_\lambda^0 p_1 - \beta S_\lambda p_0 \quad (25)$$

$$\partial_\lambda p_1 = \left[ \dot{\lambda}^{-1} \hat{L}_\lambda + \hat{A}_\lambda^1 \right] p_1 + \hat{A}_\lambda^1 p_0 + \beta S_\lambda p_0 \quad (26)$$

where we abbreviate

$$\hat{A}_\lambda^0 \equiv \hat{\Pi}_\lambda \hat{L}_\lambda^W \quad \text{and} \quad \hat{A}_\lambda^1 \equiv (1 - \hat{\Pi}_\lambda) \hat{L}_\lambda^W. \quad (27)$$

In this form, an expansion in  $\dot{\lambda}$  becomes possible. In lowest order ( $\dot{\lambda} \rightarrow 0$ ), eq. (26) implies  $\hat{L}_\lambda p_1 = 0$ . Since  $p_1$  is orthogonal to the null space of  $\hat{L}_\lambda$  by definition (24),  $p_1 = 0$  follows. For a solution of eq. (25) we explicitly calculate

$$\hat{A}_\lambda^0 f_\lambda^{\text{eq}} = -f_\lambda^{\text{eq}} \int d\mathbf{x} \frac{\partial V_\lambda}{\partial \lambda} f_\lambda^{\text{eq}} \frac{\partial}{\partial W} = -f_\lambda^{\text{eq}} \left\langle \frac{\partial V_\lambda}{\partial \lambda} \right\rangle_\lambda \frac{\partial}{\partial W}. \quad (28)$$

Using this and eq. (23) we finally obtain

$$\frac{\partial P}{\partial \lambda} = - \left\langle \frac{\partial V_\lambda}{\partial \lambda} \right\rangle_\lambda \frac{\partial P}{\partial W} \quad (29)$$

for the distribution of the work  $P(W, \lambda)$ . The solution of this equation is  $P(W) = \delta(W - \Delta F)$  with the initial condition  $P(W, 0) = \delta(W)$  following from eq. (16), where we recognize

$$\Delta F = \int_0^1 d\lambda \left\langle \frac{\partial V_\lambda}{\partial \lambda} \right\rangle_\lambda \quad (30)$$

as the change in free energy of the entire process. We thus have recovered for  $\dot{\lambda} \rightarrow 0$  the quasi-static limit as expected.

To first order in  $\dot{\lambda}$  we get from eq. (26)

$$p_1 = -\dot{\lambda} \hat{L}_\lambda^{-1} \left[ \hat{A}_\lambda^1 + \beta S_\lambda \right] p_0. \quad (31)$$

Putting this back into eq. (25) we get after a straightforward calculation a diffusion type equation for  $P(W, \lambda)$  in the form

$$\frac{\partial P}{\partial \lambda} = - \left[ \left\langle \frac{\partial V_\lambda}{\partial \lambda} \right\rangle_\lambda + \dot{\lambda} \beta \tilde{S}_\lambda \right] \frac{\partial P}{\partial W} + \dot{\lambda} \tilde{S}_\lambda \frac{\partial^2 P}{\partial W^2} \quad (32)$$

where

$$\tilde{S}_\lambda \equiv - \int d\mathbf{x} \frac{\partial V_\lambda}{\partial \lambda} \hat{L}_\lambda^{-1} S_\lambda f_\lambda^{\text{eq}}. \quad (33)$$

The solution is a Gaussian

$$P(W) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left[ -\frac{(W - \bar{W})^2}{2\sigma^2} \right] \quad (34)$$

with variance

$$\sigma^2 = 2\dot{\lambda} \int_0^1 d\lambda \tilde{S}_\lambda \quad (35)$$

and mean

$$\bar{W} = \int_0^1 d\lambda \left[ \left\langle \frac{\partial V_\lambda}{\partial \lambda} \right\rangle_\lambda + \dot{\lambda} \beta \tilde{S}_\lambda \right] = \Delta F + \frac{\beta}{2} \sigma^2. \quad (36)$$

This is the central result of the present paper [25]. First, it proves that the distribution of the work in isothermal non-equilibrium processes is Gaussian in the near-equilibrium regime. Second, we recover independently from Jarzynski's relation (3) the constraint that the mean and variance are connected according to eq. (4). Third, it yields an explicit algorithm of how to calculate these quantities.

For an assessment of the range of validity of this approximation, we recall that it is based essentially on a separation of time scales. Hence, the Gaussian distribution will be a good approximation as long as

$$\dot{\lambda}\tau \ll 1, \quad (37)$$

where  $\tau$  is an intrinsic relaxation time.

As an example, we illustrate our approach for a simple one-dimensional case, where we can compare our expansion with an exact solution [15]. We consider a colloidal particle at position  $x$  with mobility  $\mu$  trapped by an optical tweezer whose center  $y(\lambda)$  is moved at constant speed  $v$  through a viscous fluid (see FIG. 1). The potential of the trap is assumed to be harmonic near the focal point

$$V_\lambda(x) = (k/2)(x - y(\lambda))^2 \quad (38)$$

with effective strength  $k$ . In this case, the free energy is independent of  $y(\lambda)$ . For the two states, we choose with  $y(\lambda = 0) = 0$  and  $y(\lambda = 1) = L$  two positions of the trap. The switching rate becomes  $\dot{\lambda} = v/L$ , while the relaxation time is  $\tau = 1/\mu k$ .

Within our scheme, we first have to calculate  $\Psi_\lambda \equiv \hat{L}_\lambda^{-1} S_\lambda f_\lambda^{\text{eq}}$  in eq. (33) which amounts to solving the inhomogeneous differential equation

$$\hat{L}_\lambda \Psi_\lambda = S_\lambda f_\lambda^{\text{eq}}. \quad (39)$$

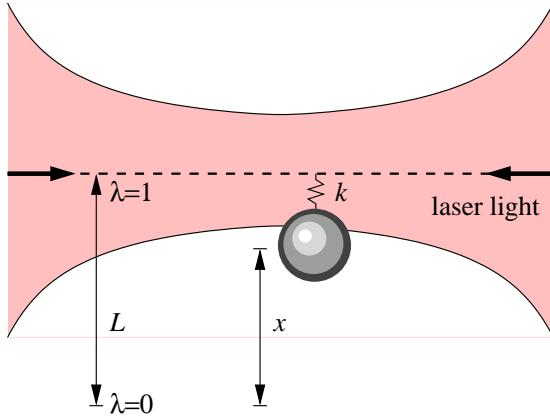


FIG. 1: Scheme of the experiment of Ref. [1]. A colloidal particle at position  $x(t)$  is dragged by an optical tweezer with focus at  $y(\lambda) = L\lambda$  through a viscous fluid. The effective potential is harmonic with a spring constant  $k$ .

With the Fokker-Planck operator from eq. (8) we get explicitly

$$\mu \frac{\partial}{\partial x} \left[ k(x - L\lambda) + \frac{1}{\beta} \frac{\partial}{\partial x} \right] \Psi_\lambda = - \sqrt{\frac{\beta k^3}{2\pi}} L(x - L\lambda) \exp \left[ -\frac{1}{2} \beta k(x - L\lambda)^2 \right] \quad (40)$$

where  $\lambda$  only appears as a parameter. This is easily solved as  $\Psi_\lambda(x) = f_\lambda^{\text{eq}}(x)Lx/\mu$  and thus the average work becomes

$$\bar{W} = L^2 k \dot{\lambda} \tau. \quad (41)$$

Of course, for this harmonic potential, the distribution  $P(W)$  is Gaussian at any driving [15]. The exact result for the mean  $\bar{W}$  as a function of  $\lambda$  reads

$$\bar{W} = L^2 k \left[ \dot{\lambda} \tau - \dot{\lambda}^2 \tau^2 \left( 1 - e^{-1/\dot{\lambda} \tau} \right) \right], \quad (42)$$

which agrees to first order in  $\dot{\lambda}$  with eq. (41) as expected.

The exact expression (42) points to an interesting property which seems not to have been discussed yet in the context of stochastic dynamics [26]. The exponentially small last term shows that the average work is non-analytic in  $\dot{\lambda}$ . We expect that if our expansion of eq. (26) was extended to the next order, some signature of this non-analyticity should show up. Therefore the approach to equilibrium (or the deviation from equilibrium) even in this almost trivial case is somewhat subtle.

In summary, we have shown for general diffusive systems that the distribution of work required to drive the system from an initial equilibrium state to another is a Gaussian for slow but finite driving. Its mean and variance can be obtained from solving an inhomogeneous differential equation involving the Fokker-Planck operator. As an exactly solvable case shows, these quantities are

non-analytic in the switching rate. This result indicates that in general calculating the next order correction to the Gaussian derived here may face fundamental difficulties.

Stimulating discussions with O. Braun and R. Finken are gratefully acknowledged as well as valuable hints and comments by H. Spohn.

---

- [1] G. M. Wang, E. M. Sevick, E. Mittag, D. J. Searles, and D. J. Evans, Phys. Rev. Lett. **89**, 050601 (2002).
- [2] D. M. Carberry, J. C. Reid, G. M. Wang, E. M. Sevick, D. J. Searles, and D. J. Evans, Phys. Rev. Lett. **92**, 140601 (2004).
- [3] J. Liphardt, S. Dumont, S. B. Smith, I. Tinoco Jr, and C. Bustamante, Science **296**, 1832 (2002).
- [4] C. Jarzynski, Phys. Rev. Lett. **78**, 2690 (1997).
- [5] C. Jarzynski, Phys. Rev. E **56**, 5018 (1997).
- [6] G. E. Crooks, Phys. Rev. E **61**, 2361 (2000).
- [7] G. Hummer and A. Szabo, Proc. Natl. Acad. Sci. U.S.A. **98**, 3658 (2001).
- [8] F. Ritort, C. Bustamante, and I. Tinoco Jr, Proc. Natl. Acad. Sci. U.S.A. **99**, 13544 (2002).
- [9] S. Park, F. Khalili-Araghi, E. Tajkhorshid, and K. Schulten, J. Chem. Phys. **119**, 3559 (2003).
- [10] S. Park and K. Schulten, J. Chem. Phys. **120**, 5946 (2004).
- [11] O. Braun, A. Hanke, and U. Seifert (2004), cond-mat/0402496.
- [12] D. M. Zuckerman and T. B. Woolf, Phys. Rev. Lett. **89**, 180602 (2002).
- [13] J. Gore, F. Ritort, and C. Bustamante, Proc. Natl. Acad. Sci. U.S.A. **100**, 12564 (2003).
- [14] M. R. Shirts, E. Bair, G. Hooker, and V. S. Pande, Phys. Rev. Lett. **91**, 140601 (2003).
- [15] O. Mazonka and C. Jarzynski (1999), cond-mat/9912121.
- [16] G. Hummer, J. Chem. Phys. **114**, 7330 (2001).
- [17] J. Hermans, J. Phys. Chem. **95**, 9029 (1991).
- [18] R. H. Wood, W. C. F. Mühlbauer, and P. T. Thompson, J. Phys. Chem. **95**, 6670 (1991).
- [19] S. R. de Groot and P. Mazur, *Non-equilibrium thermodynamics* (North-Holland, 1962).
- [20] D. A. Hendrix and C. Jarzynski, J. Chem. Phys. **114**, 5974 (2001).
- [21] C. W. Gardiner, *Handbook of Stochastic Methods* (Springer-Verlag, 1985), 2nd ed.
- [22] K. Sekimoto, Prog. Theor. Phys. Supp. **130**, 17 (1998).
- [23] N. G. van Kampen, Phys. Rep. **124**, 69 (1985).
- [24] For brevity we choose constant mobility coefficients. Otherwise the Fokker-Planck operator eq. (8) would contain spurious drift terms which would, however, not affect our main result.
- [25] Since  $S_\lambda f_\lambda^{\text{eq}}$  has no component in the null space of  $\hat{L}_\lambda$  ( $(\hat{H}_\lambda S_\lambda f_\lambda^{\text{eq}} = 0)$ ,  $\hat{L}_\lambda^{-1} S_\lambda f_\lambda^{\text{eq}}$  exists. A manifestly positive definite expression for  $\hat{S}_\lambda$  can be obtained if the distribution of the dissipated work,  $P(W - \Delta F)$ , is considered.
- [26] In quantum mechanics, exponentially small terms as in eq. (42) occur for transition probabilities of time-dependent two-state systems, see, e.g., M. V. Berry, Proc. R. Soc. A **429**, 61 (1990).